On number of turns in reduced random lattice paths

Yunjiang Jiang
Department of Mathematics
Stanford University
jyj@math.stanford.edu

Weijun Xu
Mathematical and Oxford-Man Institutes
University of Oxford
xu@maths.ox.ac.uk

September 27, 2011

Abstract

We consider the tree-reduced path of symmetric random walk on $\mathbb{Z}^d$. It is interesting to ask about the number of turns T_n in the reduced path after n steps. This question arises from inverting the signatures of lattice paths. We show that, when n is large, the mean and variance of T_n in the asymptotic expansion have the same order as n, while the lower terms are $O(1)$. We also obtain limit theorems for T_n, including large deviations principle, central limit theorem, and invariance principle. Similar techniques apply to other finite patterns in a lattice path.

Key words: signature of a path, reduced word, number of turns

Mathematics Subject Classification: 60

1 Introduction

Let G be the free group with d generators $e_1, \cdots, e_d$. Start with the empty word at time 0. At each time k, choose one from the $2d$ elements (d generators and their inverses) uniformly randomly to multiply the current word on the right. For example, the first six choices:

$$e_2, e_3, {e_3}^{-1}, e_2, {e_1}^{-1}, e_4$$

will produce the reduced word $e_2 e_2 {e_1}^{-1} e_4$ at time 6. Every word at time n has a unique reduced word with length at most n. It is then interesting to ask about the length and number of turns in the reduced word.

Definition 1.1. *Let w be a word, and $\hat{w} = a_{i_1} \cdots a_{i_k}$ be its reduced word, where a_j is either e_j or ${e_j}^{-1}$. Define the number of turns of w to be $T_n = \#\{a_{i_j} a_{i_{j+1}} : a_{i_j} \neq a_{i_{j+1}}, 1 \leq j \leq k-1\}$.*

Then, two words have the same number of turns if they reduce to the same word. In the above example, the number of turns in the reduced word $e_2 e_2 {e_1}^{-1} e_4$ is 2. In another language, it is the number of times a reduced path of random walk has switched its direction.

The main goal of this paper is to calculate asymptotics for T_n when n is large. The question of estimating T_n arises from inverting signature for lattice paths, where at most $T+2$ terms in the signature are needed for inversion if one knows in advance that the reduced lattic path has T turns.

1.1 Motivation from inversion of signature for axis paths

In this subsection, we give some background material on path-signature that motivates our study of the current problem. A path $\gamma : [s,t] \to \mathbb{R}^d$ is a continuous function mapping a time interval into $\mathbb{R}^d$. The length of the path is defined as

$$|\gamma| := \sup_{\mathcal{P}} d(\gamma(t_i), \gamma(t_{i+1})),$$

where d the metric on $\mathbb{R}^d$, and the supremum is taken over all finite partitions of $[s,t]$. If $|\gamma| < +\infty$, we say γ has bounded variation. Let $BV(\mathbb{R}^d)$ denote the space of all paths of bounded variations in $\mathbb{R}^d$.

Definition 1.2. *Let $\gamma : [s,t] \to \mathbb{R}^d$ be an element in $BV(\mathbb{R}^d)$. The signature of γ, $X_{s,t}(\gamma)$, is defined as:*

$$X_{s,t}(\gamma) = 1 + X^1_{s,t}(\gamma) + \cdots + X^n_{s,t}(\gamma) + \cdots ,$$

where

$$X^n_{s,t}(\gamma) = \int_{s<u_1 \cdots <u_n<t} d\gamma(u_1) \otimes \cdots \otimes d\gamma(u_n) \tag{1}$$

as an element in $(\mathbb{R}^d)^{\otimes n}$.

Let $(e_1, e_2, \cdots, e_d)$ be a standard basis of $\mathbb{R}^d$, then γ can be written as $(\gamma_1, \gamma_2, \cdots, \gamma_d)$. If $w = e_{i_1} \cdots e_{i_n}$ is a word of length n, we write

$$C_{s,t}(w) = C_{s,t}(e_{i_1} \cdots e_{i_n}) = \int_{s<u_1<\cdots<u_n<t} d\gamma_{i_1}(u_1) \cdots d\gamma_{i_n}(u_n)$$

as the coefficient of w. As all words of length n form a basis of $V^{\otimes n}$, we can rewrite $X^n_{s,t}(\gamma)$ as the linear combination of basis elements:

$$X^n_{s,t}(\gamma) = \sum_{|w|=n} C_{s,t}(w) w, \tag{2}$$

where the sum is taken over all words of length n.

Re-parametrizing the path does not change the signature. For any path $\alpha : [0,s] \to V$ and $\beta : [0,t] \to V$, we can form the concatenation $\alpha * \beta : [0, s+t] \to V$, as follows:

$$\alpha * \beta(u) := \begin{cases} \alpha(u), u \in [0,s] \\ \beta(u-s) + \alpha(s) - \alpha(0), u \in [s, s+t] \end{cases},$$

and similarly, the decomposition of one path into two can be carried out in the same fashion.

For any path $\gamma : [s,t] \to V$, the path "γ run backwards", γ^{-1}, is defined as:

$$\gamma^{-1}(u) := \gamma(s+t-u), u \in [s,t],$$

and the trajectories of $\gamma * \gamma^{-1}$ cancel out each other.

Concatenation and "backwards" of paths of bounded variation are still paths of bounded variation. In fact, we have $|\alpha * \beta| = |\alpha| + |\beta|$, and $|\gamma^{-1}| = |\gamma|$. The following proposition, first proved by Chen ([1]), asserts that the signature map is a homomorphism from $BV(\mathbb{R}^d)$ to the tensor algebra.

Proposition 1.3. *Let $\alpha, \beta \in BV^(\mathbb{R}^d)$. Then, $X(\alpha * \beta) = X(\alpha) \otimes X(\beta)$.*

Hambly and Lyons ([4]) showed that if $\alpha, \beta \in BV(\mathbb{R}^d)$, then $X(\alpha) = X(\beta)$ if and only if $\alpha * \beta^{-1}$ is tree-like, a continuous analogue of a null path. This tree-like relation defines an equivalence relation on $BV(\mathbb{R}^d)$. Within every equivalent class, there is a unique path with minimal length, called the tree reduced path. An interesting question would be, given a signature X of some path of bounded variation, can one reconstruct the tree reduced path with the same signature X?

For the case of axis paths, the answer was provided by Lyons and Xu in [6].

Definition 1.4. *$\gamma : [s,t] \to \mathbb{R}^d$ is a (finite) axis path if its movements are parallel to the Euclidean coordinate axes, has finitely many turns, and each straight line component has finite length.*

Any axis path has a unique reduced axis path; integer lattice paths are special cases of axis paths. An $\mathbb{R}^d$ axis path can move in d different directions (up to the sign). At time 0, it starts to move along a direction e_{i_1} for some distance r_1; then it turns a right angle, and moves along e_{i_2} for a distance r_2, and so on, and stops after finitely many turns. Thus, up to re-parametrization, an axis path γ can be represented as:

$$\gamma = (r_1 e_{i_1}) * \cdots * (r_n e_{i_n}) \tag{3}$$

where r_i's are real numbers, with the sign denoting the direction[1].

Using Chen's identity (Proposition 1.3), the signature of γ can be conveniently expressed as

$$X(\gamma) = \exp(r_1 e_{i_1}) \otimes \cdots \otimes \exp(r_n e_{i_n}),$$

which should be understood as the product of n power series in the letters $\{e_{i_1}, \cdots, e_{i_n}\}$.

If γ is already in its reduced form, then it is clear that $i_k \neq i_{k+1}$, and we call the word $w = (e_{i_1}, \cdots, e_{i_n})$ the shape of γ. If a word w is in its reduced form, we use $|w|$ to denote the number of letters in w, or the length of w. We introduce the notion of square free words to characterize an axis path.

Definition 1.5. *Let $w = e_{i_1} \cdots e_{i_n}$ be a word. We call it a square free word if $\forall k \leq n-1$, $i_k \neq i_{k+1}$.*

The following theorem, provided by Lyons and Xu ([6]), gives an inversion procedure for finite axis paths.

Theorem 1.6. *For any finite axis path γ, there exists a unique square free word w with the property that $C(w) \neq 0$, and that if w' is any other square free word with $C(w') \neq 0$, then $|w'| < |w|$. Furthermore, suppose the unique longest square free word is $w = e_{i_1} \cdots e_{i_n}$, and let*

$$w_k := e_{i_1} \cdots e_{i_{k-1}} e_{i_k}^2 \cdot e_{i_{k+1}} \cdot e_{i_n},$$

which has length $n+1$, then we have

$$\gamma = (r_1 e_{i_1}) * \cdots * (r_n e_{i_n}),$$

where $r_k = \frac{2C(w_k)}{C(w)}$.

Thus, one sees if an axis path has n turns, then at most $n+2$ terms in the signature are needed for inversion. For a lattice path with length L, it can have at most $L-1$ turns, so we only need the first $L+1$ terms in the signature to recover it.

In pratice, lattice paths are often generated by drawing n letters and their inverses uniformly randomly from an alphabet, and putting them in the order they are drawn. It is then intersting to ask about the number of turns in its reduced path.

1.2 Outline of the method and summary of results

If one writes $T_n = \sum_{i=1}^n V_i$, where V_i denote the number of turns created at step i. In general, V_i can be 1, 0, or -1, and are correlated. The distribution of V_i depends on the whole history in the past.

On the other hand, one can condition on the length of reduced path L_n. Then $T_n|L_n$ has a binomial distribution. A detailed study of L_n yields asymptotic behaviors of T_n. A natural coupling $L_n = S_n + D_n$ simplifies the study of L_n, where S_n is a sum of n i.i.d.'s, and D_n is dominated by a geometric random variable.

The main results in this paper are:

[1] We mean $-re_j = re_j^{-1}$.

Proposition 2.5. $\mathbb{E}T_n - \frac{2(d-1)^2}{d(2d-1)}n \to -\frac{2d^2-4d+1}{d(2d-1)}$, var $T_n - \frac{2(d-1)^2(5d-2)}{d^2(2d-1)^2}n$ *also converges.*

Theorem 3.1. (Large Deviations Principle) *The sequence of the laws for the random variables* $\{\frac{T_n}{n}\}_{n\geq 1}$ *satisfy the large deviations principle with rate function*

$$I(x) = \sup_\theta[\theta x - \log h(\theta)],$$

where

$$h(\theta) = \begin{cases} \frac{1}{2d}[2(d-1)e^\theta + \frac{2d-1}{1+2(d-1)e^\theta} + 1], \theta \geq \log\frac{\sqrt{2d-1}-1}{2(d-1)} \\ \frac{\sqrt{2d-1}}{d}, \qquad \theta < \log\frac{\sqrt{2d-1}-1}{2(d-1)} \end{cases},$$

Theorem 4.5. (Invariance Principle) *For each n,define a* $C^0([0,1])$*-valued random variable* $\{W_t^{(n)} : t \in [0,1]\}$ *by*

$$W_t^{(n)} = \frac{1}{\sigma\sqrt{n}}[T_{tn} - \frac{2(d-1)^2}{d(2d-1)}tn]$$

for $t \in \frac{1}{n}[n]$*, and linearly interpolated for other values of* t*. Then the sequence converges in law to the standard one dimensional Brownian motion on* $[0,1]$ *as* $n \to \infty$.

As a generalization, analogous results hold for the number of occurrences of any finite collection of finite length pattern $\mathcal{P} = \{P_i = (e_{i_1}, \ldots, e_{i_{k_i}})\}$ in a lattice path: the key is to establish central limit theorem for the number of occurences of elements in $\mathcal{P}$ conditioned on the length of the path L_t, which is essentially an i.i.d. sum of m-dependent random variables (see [7]), where m is bounded above by $\max_i k_i$. The number of turns T_t corresponds to $\mathcal{P} = \{P_{ij} = (e_i, e_j) : i \neq j\}$. For the sake of clarity, we will focus only on the number of turns.

The paper is organized as follows:

In section 2, we show that the lower order terms in the mean and variance of L_n are $O(1)$; we then obtain similar results for T_n. A key ingradient in the derivation is to prove cov (S_n, D_n) is $O(1)$. We compare it with cov (S_{n+1}, D_{n+1}), and show that their differnce decays exponentially with n, thus proving convergence.

Section 3 is devoted to the proof of the large deviations principle for $\{\frac{T_n}{n}\}$. We derive the rate function, and thus prove the principle, by comparison of the Laplace transform of L_n with that of S_n. It turns out that the rate function deviates from the normal one as predicted by S_n on the lower side of the real line.

In section 4, we prove the central limit theorem and invariance principle for T_n. This result shows that, although the components of T_n are correlated, the increments are still asymptotically independent under proper scaling.

2 Lower order terms in mean and variance

Let L_n denote the length of the reduced path after n steps. Then, $T_n|L_n$ has a binomial distribution with parameters $(L_n - 1, \frac{2d-2}{2d-1})$. Let X_i be a sequence of indenpedent and indentically distributed random variables with $\mathbb{P}(X_i = 1) = \frac{2d-1}{2d}$, and $\mathbb{P}(X_i = -1) = \frac{1}{2d}$. Let $L_0 = 0$, then L_n can be defined inductively as follows:

$$L_{i+1} = L_i + X_{i+1}, L_i > 0$$
$$L_{i+1} = 1, L_i = 0$$

We want to compare $L_n - S_n$, where $S_n = \sum_{i=1}^n X_i$. It is well known that $\frac{L_n}{n} \to \frac{d-1}{d}$ almost surely. We compute a finer estimate to show that $\mathbb{E}L_n - \frac{d-1}{d}n = O(1)$.

Since $L_i - S_i$ does not change when L is away from 0, so difference only occurs when L hits 0. Let R_n denote the number of times that L_n hits to 0 after the first step up to time n. Since $\mathbb{P}(L$ ever comes back

to 0)$=\frac{1}{2d-1}$, R_n converges to a geometric distributed random varible R, with $\mathbb{P}(R=k) = \frac{2d-2}{2d-1}(\frac{1}{2d-1})^{k-1}$ (see e.g. [3]). Here, step 0 is counted as a return, because L and S can be different in the first move. Let $D_n = L_n - S_n$, then $\frac{1}{2}D_n|R_{n-1}$ has a binomial distribution with parameters $(R_{n-1}, \frac{1}{2d})$. In particular, $D_n \le 2R_{n-1}$. So the mean of the difference is:

$$\mathbb{E}D_n = \mathbb{E}\mathbb{E}(D_n|R_{n-1}) = \frac{1}{d}\mathbb{E}R_{n-1} \to \frac{2d-1}{2d(d-1)}$$

Thus, we get an error term for $\mathbb{E}L_n$:

Lemma 2.1. *Let L_n denote the length of the reduced word after n steps, then*

$$\lim_{n\to\infty}(\mathbb{E}L_n - \frac{d-1}{d}n) = \frac{2d-1}{2d(d-1)}.$$

Now to compute the lower order terms in var L_n. Since var L_n = var S_n + 2cov (S_n, D_n) + var D_n, it suffices to show that cov $(S_n, D_n) = O(1)$. We show it in the following lemma.

Lemma 2.2. *Under the above coupling $L_n = S_n + D_n$, we have:*

$$\lim_{n\to\infty}(\mathbb{E}S_nD_n - \mathbb{E}S_n\mathbb{E}D_n) = -u(d),$$

where $0 \le u(d) \le \frac{2d^2(2d-1)(d^2+2d-1)}{(d-1)^5}$.

Proof. Let $U_n = (S_n - \frac{d-1}{d}n)D_n$. We show that $\mathbb{E}U_{n+1} - \mathbb{E}U_n$ decays exponentially fast.

$$\begin{aligned} U_{n+1} &= 1_{\{L_n=0\}}(S_n + X_{n+1} - \frac{d-1}{d}(n+1))(D_n + 1 - X_{n+1}) + 1_{\{L_n>0\}}(S_n + X_{n+1} - \frac{d-1}{d}(n+1))D_n \\ &= U_n + (X_{n+1} - \frac{d-1}{d})D_n + 1_{\{L_n=0\}}(1 - X_{n+1})(S_{n+1} - \frac{d-1}{d}(n+1)) \end{aligned}$$

Since X_{n+1} is independent of D_n, and $-n \le S_n \le 0$ when conditioned on $L_n = 0$, taking expectation on both sides yields:

$$-4(n+1)\mathbb{P}(L_n = 0) \le \mathbb{E}U_{n+1} - \mathbb{E}U_n \le 0.$$

Since $\mathbb{P}(L_{2n+1} = 0) = 0$, and

$$\begin{aligned} \mathbb{P}(L_{2n} = 0) &\le \mathbb{P}(S_{2n} \le 0) \\ &\le \sum_{k=0}^{n} 2^{2n}(\frac{1}{2d})^{n+k}(\frac{2d-1}{2d})^{n-k} \\ &\le \frac{2d-1}{2(d-1)}(\frac{2d-1}{d^2})^n \end{aligned}$$

This shows that $\mathbb{E}U_n$ is decreasing and bounded below, and thus it has a finite limit.
Adding up all $(\mathbb{E}U_n - \mathbb{E}U_{n-1})$ gives $\mathbb{E}U_n \to -u(d)$, where $0 \le u(d) \le \frac{2d^2(2d-1)(d^2+2d-1)}{(d-1)^5}$. □

Remark 2.3. *The negative correlation agrees with one's probabilistic intuition: when S_n is small, the process L_n tends to visit 0 more times, and thus D_n is likely to be large.*

Proposition 2.4. *Let $u(d)$ be the constant as in the previous lemma. Then,*

$$\lim_{n\to\infty}(var\ L_n - \frac{2d-1}{d^2}n) = \beta(d),$$

where $\beta(d) = -2u(d) + \frac{(2d-1)(4d^2-6d+3)}{4d^2(d-1)^2}$.

Proof. Since var $D_n = \mathbb{E}\text{var }(D_n|R_{n-1}) + \text{var }\mathbb{E}(D_n|R_{n-1}) \to \frac{(2d-1)(4d^2-6d+3)}{4d^2(d-1)^2}$, we have:

$$\text{var } L_n - \frac{2d-1}{d^2}n = 2\text{cov }(S_n, D_n) + \text{var } D_n \to \beta(d),$$

where $\beta(d) = -2u(d) + \frac{(2d-1)(4d^2-6d+3)}{4d^2(d-1)^2}$. □

Combining the above estimates for L_n, we then have similar estimates for T_n:

$$\mathbb{E}T_n = \mathbb{E}\mathbb{E}(T_n|L_n) = \frac{2d-2}{2d-1}\mathbb{E}L_n - \frac{2d-2}{2d-1}$$

$$\begin{aligned}\text{var } T_n &= \mathbb{E}\text{var }(T_n|L_n) + \text{var }\mathbb{E}(T_n|L_n) \\ &= \frac{2(d-1)}{(2d-1)^2}\mathbb{E}L_n + \frac{4(d-1)^2}{(2d-1)^2}\text{var } L_n - \frac{2(d-1)}{(2d-1)^2},\end{aligned}$$

which gives the following proposition:

Proposition 2.5. *Let $\beta(d)$ be the error term in var L_n as above. Then,*

$$\lim_{n\to\infty}(\mathbb{E}T_n - \frac{2(d-1)^2}{d(2d-1)}n) = -\frac{2d^2-4d+1}{d(2d-1)},$$

and

$$\lim_{n\to\infty}(\textit{var } T_n - \frac{2(d-1)^2(5d-2)}{d^2(2d-1)^2}n) = \frac{4(d-1)^2}{(2d-1)^2}\beta(d) - \frac{2d^2-4d+1}{d(2d-1)^2}.$$

3 Large deviations

The goal of this section is to prove the following large deviations theorem for $\frac{T_n}{n}$.

Theorem 3.1. *The sequence of the laws for the random variables $\{\frac{T_n}{n}\}_{n\geq 1}$ satisfy the large deviations principle with rate function*

$$I(x) = \sup_\theta[\theta x - \log h(\theta)],$$

where

$$h(\theta) = \begin{cases} \frac{1}{2d}[2(d-1)e^\theta + \frac{2d-1}{1+2(d-1)e^\theta} + 1], & \theta \geq \log\frac{\sqrt{2d-1}-1}{2(d-1)} \\ \frac{\sqrt{2d-1}}{d}, & \theta < \log\frac{\sqrt{2d-1}-1}{2(d-1)} \end{cases},$$

We postpone the proof of this theorem to the end of the section. In light of Gartner-Ellis theorem (see [2] section 2.3), it suffices to show that

$$\lim_{n\to+\infty}(\mathbb{E}e^{\theta T_n})^{\frac{1}{n}} = h(\theta)$$

for every $\theta \in \mathbb{R}$, and the limit h is essentially smooth.

Note that $T_n|L_n$ has a binomial distribution with parameter $(\frac{2d-2}{2d-1}, L_n - 1)$ for $L_n \geq 1$, and $T_n = 0$ if $L_n = 0$, we have

$$\mathbb{E}e^{\theta T_n} = \mathbb{P}(L_n = 0) + \mathbb{P}(L_n > 0)\mathbb{E}w(\theta)^{L_n - 1}, \tag{4}$$

where $w(\theta) = \frac{2d-2}{2d-1}e^\theta + 1$. The first term is bounded by

$$\begin{aligned}\mathbb{P}(L_n = 0) &\leq \sum_{k=0}^{n} \mathbb{P}(S_n = -k) \\ &\leq C(\frac{\sqrt{2d-1}}{d})^n,\end{aligned}$$

and we need to estimate $\mathbb{E}w(\theta)^{L_n}$ for large n.

In the context below, we regard w to be a positive real number independent of θ, and study the asymptotics of $(\mathbb{E}w^{L_n})^{\frac{1}{n}}$ as $n \to +\infty$.

Proposition 3.2. *If $w \geq 1$, then we have*

$$\lim_{n\to+\infty} (\mathbb{E}w^{L_n})^{\frac{1}{n}} = \frac{2d-1}{2d}w + \frac{1}{2d} \cdot \frac{1}{w}.$$

Proof. We compare the difference between $\mathbb{E}w^{L_{n+1}}$ and $\mathbb{E}w^{L_n}$:

$$\begin{aligned}\mathbb{E}w^{L_{n+1}} &= \mathbb{E}1_{\{L_n=0\}}w^{L_{n+1}} + \mathbb{E}1_{\{L_n>0\}}w^{L_{n+1}} \\ &= w\mathbb{P}(L_n = 0) + \mathbb{E}1_{\{L_n>0\}}w^{L_n+X_{n+1}} \\ &= w\mathbb{P}(L_n = 0) + \mathbb{E}w^{X_{n+1}}\mathbb{E}1_{\{L_n>0\}}w^{L_n} \\ &= w\mathbb{P}(L_n = 0) + \mathbb{E}w^{X_{n+1}}\mathbb{E}w^{L_n} - \mathbb{E}w^{X_{n+1}}\mathbb{P}(L_n = 0) \\ &= (\frac{2d-1}{2d}w + \frac{1}{2d}\frac{1}{w})\mathbb{E}w^{L_n} + \frac{1}{2d}(w - \frac{1}{w})\mathbb{P}(L_n = 0)\end{aligned}$$

Let $x_n = \mathbb{E}w^{L_n}, a = \frac{2d-1}{2d}w + \frac{1}{2d}\frac{1}{w}, b = \frac{1}{2d}(w - \frac{1}{w}), p_n = \mathbb{P}(L_n = 0)$, we have the following recursive relation:

$$x_n = ax_{n-1} + bp_{n-1}$$

Since $x_1 = w$, adding them up yields:

$$x_n = a^{n-1}w + b(a^{n-2}p_1 + a^{n-3}p_2 + \cdots + ap_{n-2} + p_{n-1})$$

For $w \geq 1$, we have $b \geq 0$. In this case, since $a^{n-1} \leq x_n \leq na^{n-1}$ from the expression above, we get:

$$\lim_{n\to\infty} (x_n)^{\frac{1}{n}} = a = \frac{2d-1}{2d}w + \frac{1}{2d}\frac{1}{w},$$

thus proving the proposition. ∎

The situation for $w \in (0, 1)$ is more involved. We prove it based on comparison with $(\mathbb{E}w^{S_n})^{\frac{1}{n}}$. Note that S_n is a sum of i.i.d., so by Cramer's theorem, it satisfies large deviations principle with rate function

$$J(x) = \sup_\theta[\theta x - \log(\frac{2d-1}{2d}e^\theta + \frac{1}{2d}e^{-\theta})].$$

Lemma 3.3. *For any $w > 0$, the equation*

$$w^\alpha e^{-J(\alpha)} = \frac{2d-1}{2d}w + \frac{1}{2d} \cdot \frac{1}{w} \tag{5}$$

has a unique solution at $\alpha^* = \frac{(2d-1)w^2-1}{(2d-1)w^2+1}$. Furthermore, α^* is the global maximizer for

$$f_w(\alpha) = w^\alpha e^{-J(\alpha)}.$$

Proof. We first give an expression of J in terms of α only. It is clear that $J(\alpha) = +\infty$ for $|\alpha| > 1$. For $\alpha \in [-1, 1]$, the maximizer θ^* is

$$\theta^*(\alpha) = \frac{1}{2}[\log\frac{1+\alpha}{1-\alpha} - \log(2d-1)], \alpha \in [-1, 1],$$

passing to the limit $\pm\infty$ for $\alpha = \pm 1$. Substituting into J, we have

$$J(\alpha) = \frac{1}{2}\alpha[\log\frac{1+\alpha}{1-\alpha} - \log(2d-1)] - \log\frac{\sqrt{2d-1}}{2d} - \log(\sqrt{\frac{1+\alpha}{1-\alpha}} + \sqrt{\frac{1-\alpha}{1+\alpha}})$$

for $\alpha \in [-1, 1]$. Differentiating with respect to α, we obtain

$$J'(\alpha) = \frac{1}{2}[\log\frac{1+\alpha}{1-\alpha} - \log(2d-1)] \tag{6}$$

for $\alpha \in (-1, 1)$. Note that

$$f'_w(\alpha) = \frac{d}{d\alpha}(w^\alpha e^{-J(\alpha)}) = w^\alpha e^{-J(\alpha)}(\log w - J'(\alpha)),$$

and since J is convex, f_w has the global maximizer α^* satisfying

$$J'(\alpha^*) = \log w.$$

By (6), solving the above first order condition yieds

$$\alpha^* = \frac{(2d-1)w^2 - 1}{(2d-1)w^2 + 1},$$

and thus

$$f_w(\alpha^*) = \frac{2d-1}{2d}w + \frac{1}{2d} \cdot \frac{1}{w}.$$

Since α^* is the global maximier of f_w, we conclude that equation (5) has a unique solution at α^*. $\square$

Proposition 3.4. *Let* $\alpha^* = \alpha^*(w) = \frac{(2d-1)w^2-1}{(2d-1)w^2+1}$ *be as above, then*

$$\lim_{\epsilon \downarrow 0} \lim_{n \to +\infty} (\mathbb{E}w^{S_n} 1_{\{\frac{S_n}{n} \in (\alpha^* - \epsilon, \alpha^* + \epsilon)\}}) = \frac{2d-1}{2d}w + \frac{1}{2d} \cdot \frac{1}{w}.$$

This proposition shows that the major contribution for $\mathbb{E}w^{S_n}$ are from the S_n's with values near $\alpha^* n$.

Lemma 3.5. *(a) Let* $\alpha > 0$. $\forall \epsilon \in (0, \alpha)$, $\forall \delta > 0$, $\exists N = N(\alpha, \epsilon, \delta)$ *such that*

$$\mathbb{P}(L_n - S_n \le \delta n | \frac{S_n}{n} \in (\alpha - \epsilon, \alpha + \epsilon)) \ge \frac{1}{2}$$

for all $n \ge N$.
(b) $\forall \epsilon, \delta > 0$, $\exists N = N(\epsilon, \delta)$ *such that*

$$\mathbb{P}(L_n \le \delta n | S_n \le -\epsilon n) \ge \frac{1}{2}$$

for all $n \ge N$.

Proof. We prove part(a), and the proof for part (b) is similar. Observe that

$$D_n = L_n - S_n = 2|\min_{0 \le k \le n} S_k|,$$

we first consider the quantity $\mathbb{P}(\min_{1 \le k \le n} S_k \ge -\delta n | \frac{S_n}{n} = \alpha n)$, where without loss of generality, we have assumed $\frac{1+\alpha}{2}n$ is an integer, and have replaced $\frac{\delta}{2}$ by δ. Once conditioned on the event $\{\frac{S_n}{n} = \alpha n\}$, all possible paths contain $\binom{n}{\frac{1+\alpha}{2}n}$ positive movements and $\binom{n}{\frac{1-\alpha}{2}n}$ negative movements. Since all these

paths have the same (conditional) weight, the quantity $\mathbb{P}(\min_{1\le k\le n} S_k \ge -\delta n | \frac{S_n}{n} = \alpha n)$ is independent of d. Thus, we may assume $d = 1$, where all paths are the trajectories of the (conditional) simple symmetric random walk. That is,

$$\mathbb{P}(\min_{0\le k\le n} S_k < -\delta n | S_n = \alpha n) = \mathbb{P}(\min_{0\le k\le n} \tilde{S}_k < -\delta n | \tilde{S}_n = \alpha n),$$

where $\tilde{S}_n$ is a one-dimensional simple symmetric random walk. By reflection principle, we have

$$\mathbb{P}(\min_{0\le k\le n} \tilde{S}_k < -\delta n, \tilde{S}_n = \alpha n) = \mathbb{P}(\tilde{S}_n = -\lfloor(\alpha + 2\delta)\rfloor n).$$

Using Stirling's approximation, we estimate the ratio

$$\begin{aligned}
\frac{\mathbb{P}(\tilde{S}_n = -\lfloor(\alpha+2\delta)n)\rfloor}{\mathbb{P}(\tilde{S}_n = \alpha n)} &= \binom{n}{\frac{1-\alpha-2\delta}{2}n} / \binom{n}{\frac{1+\alpha}{2}n} \\
&= (\frac{1-\alpha}{2}n)!(\frac{1+\alpha}{2}n)!/[(\frac{1-\alpha-2\delta}{2}n)!(\frac{1+\alpha+2\delta}{2}n)!] \\
&\approx [\frac{(1-\alpha)^{1-\alpha}(1+\alpha)^{1+\alpha}}{(1-\alpha-2\delta)^{1-\alpha-2\delta}(1+\alpha+2\delta)^{1+\alpha+2\delta}}]^{\frac{n}{2}}.
\end{aligned}$$

It is straightforward to check that $(1-x)\log(1-x) + (1+x)\log(1+x)$ is increasing in $(0,1)$, thus the quantity in the bracket in the last line is strictly less than one, and the ratio decays exponentially. So, we have

$$\begin{aligned}
&\mathbb{P}(\min_{0\le k\le n} S_k < -\delta n | \frac{S_n}{n} \in (\alpha-\epsilon, \alpha+\epsilon)) \\
&= \sum_{\beta\in(\alpha-\epsilon,\alpha+\epsilon)} \mathbb{P}(\min_{1\le k\le n} \tilde{S}_k < -\delta n, \frac{\tilde{S}_n}{n} \in (\alpha-\epsilon,\alpha+\epsilon))/\mathbb{P}(\frac{\tilde{S}_n}{n} \in (\alpha-\epsilon,\alpha+\epsilon)) \\
&\le n \sup_{\beta\in(\alpha-\epsilon,\alpha+\epsilon)} \mathbb{P}(\tilde{S}_n = -(\beta+2\delta)n)/\mathbb{P}(\frac{\tilde{S}_n}{n} \in (\alpha-\epsilon,\alpha+\epsilon)) \\
&\to 0,
\end{aligned}$$

and consequently

$$\mathbb{P}(\min_{0\le k\le n} \ge -\delta n | \frac{S_n}{n} \in (\alpha-\epsilon,\alpha+\epsilon)) \to 1$$

as $n \to +\infty$, thus proving (a). The proof for part (b) is essentially the same, and we omit it here. $\square$

Proposition 3.6. *For $w \in (0,1)$, the limit $(\mathbb{E}w^{L_n})^{\frac{1}{n}}$ exists as $n \to +\infty$, and we have*

$$\lim_{n\to+\infty} (\mathbb{E}w^{L_n})^{\frac{1}{n}} = \begin{cases} \frac{2d-1}{2d}w + \frac{1}{2d}\cdot\frac{1}{w}, & w \in (\frac{1}{\sqrt{2d-1}}, 1) \\ \frac{\sqrt{2d-1}}{d}, & w \in (0, \frac{1}{\sqrt{2d-1}}] \end{cases},$$

Proof. 1. $w \in (\frac{1}{\sqrt{2d-1}}, 1)$, $\alpha^* = \frac{(2d-1)w^2-1}{(2d-1)w^2+1} > 0$.

In this case, we have

$$\begin{aligned}
(\mathbb{E}w^{L_n})^{\frac{1}{n}} &\ge [\mathbb{E}(w^{L_n}1_{\{L_n-S_n\le\delta n\}}|\frac{S_n}{n} \in (\alpha^*-\epsilon,\alpha^*+\epsilon))\mathbb{P}(\frac{S_n}{n} \in (\alpha^*-\epsilon,\alpha^*+\epsilon))]^{\frac{1}{n}} \\
&\ge w^{\alpha^*+\epsilon+\delta}\mathbb{P}(\frac{S_n}{n} \in (\alpha^*-\epsilon,\alpha^*+\epsilon))^{\frac{1}{n}}\mathbb{P}(L_n - S_n \le \delta n|\frac{S_n}{n} \in (\alpha^*-\epsilon,\alpha^*+\epsilon))^{\frac{1}{n}}.
\end{aligned}$$

Taking $n \to +\infty$ on both sides yields

$$\liminf_{n\to+\infty} (\mathbb{E}w^{L_n})^{\frac{1}{n}} \ge w^{\alpha^*+\epsilon+\delta} \cdot e^{-I(\alpha^*+\epsilon)}.$$

Since ϵ and δ are arbitrary, we have

$$\liminf_{n\to+\infty} (\mathbb{E}w^{L_n})^{\frac{1}{n}} \ge w^{\alpha^*}e^{-I(\alpha^*)} = \frac{2d-1}{2d}w + \frac{1}{2d}\cdot\frac{1}{w}.$$

On the other hand, $\mathbb{E}w^{L_n} \leq \mathbb{E}w^{S_n}$ for all $w < 1$, so

$$\limsup_{n\to+\infty}(\mathbb{E}w^{L_n})^{\frac{1}{n}} \leq \frac{2d-1}{2d}w + \frac{1}{2d}\cdot\frac{1}{w}.$$

Thus, we conclude that

$$\lim_{n\to+\infty}(\mathbb{E}w^{L_n})^{\frac{1}{n}} = \frac{2d-1}{2d}w + \frac{1}{2d}\cdot\frac{1}{w}$$

for $w \in (\frac{\sqrt{2d-1}}{d}, 1)$.

2. $w \in (0, \frac{1}{\sqrt{2d-1}})$, and $\alpha^* < 0$.
   In this case, we have

$$\begin{aligned}(\mathbb{E}w^{L_n})^{\frac{1}{n}} &\geq [\mathbb{E}(w^{L_n}1_{\{L_n\leq\delta n\}}|S_n \leq -\epsilon n)\mathbb{P}(S_n \leq -\epsilon n)]^{\frac{1}{n}} \\ &\geq w^{\delta}\mathbb{P}(L_n \leq \delta n|S_n \leq -\epsilon n)^{\frac{1}{n}}\mathbb{P}(S_n \leq -\epsilon n)^{\frac{1}{n}}.\end{aligned}$$

   Again, sending $n \to +\infty$ and take ϵ, δ arbitrarily small yields

$$\liminf_{n\to+\infty}(\mathbb{E}w^{L_n})^{\frac{1}{n}} \geq e^{-I(0)} = \frac{\sqrt{2d-1}}{d}.$$

   On the other hand,

$$\lim_{w\downarrow\frac{1}{\sqrt{2d-1}}}\lim_{n\to+\infty}(\mathbb{E}w^{L_n})^{\frac{1}{n}} = \frac{\sqrt{2d-1}}{d},$$

   by monotonocity, we have

$$\limsup_{n\to+\infty}(\mathbb{E}w^{L_n})^{\frac{1}{n}} \leq \frac{\sqrt{2d-1}}{d},$$

   and thus

$$\limsup_{n\to+\infty}(\mathbb{E}w^{L_n})^{\frac{1}{n}} = \frac{\sqrt{2d-1}}{d}$$

   for all $w \in (0, \frac{1}{\sqrt{2d-1}})$.

3. $w = \frac{1}{\sqrt{2d-1}}$, $\alpha^* = 0$. Again, by monotonocity in w, we have

$$\lim_{n\to+\infty}(\mathbb{E}\sqrt{2d-1}^{-L_n})^{\frac{1}{n}} = \frac{\sqrt{2d-1}}{d}.$$

□

The following corollary is an immediate consequence of Propositions 3.2 and 3.6.

Corollary 3.7. *The laws for the random variables $\{\frac{L_n}{n}\}_{n\geq1}$ satisfy the large deviations principle with rate function*

$$I_L(x) = \sup_{\theta}[\theta x - \log h_L(\theta)],$$

where

$$h_L(\theta) = \begin{cases} \frac{2d-1}{2d}e^{\theta} + \frac{1}{2d}e^{-\theta}, & \theta \geq -\frac{1}{2}\log(2d-1) \\ \frac{\sqrt{2d-1}}{d}, & \theta < \log\frac{\sqrt{2d-1}-1}{2(d-1)} \end{cases},$$

Now we are in a position to prove Theorem 3.1.

Proof of Theorem 3.1.

Proof. Recall that

$$\mathbb{E}e^{\theta T_n} = \mathbb{P}(L_n = 0) + \mathbb{P}(L_n \geq 1)\mathbb{E}w(\theta)^{L_n - 1},$$

where we have

$$\mathbb{P}(L_n = 0) \leq C(\frac{\sqrt{2d-1}}{d})^n$$

for all n. On the other hand, Propositions 3.2 and 3.6 imply that

$$\lim_{n\to+\infty} (\mathbb{E}w^{L_n})^{\frac{1}{n}} \geq \frac{\sqrt{2d-1}}{d}$$

for all $w > 0$. Thus, we see that

$$\lim_{n\to+\infty} (\mathbb{E}e^{\theta T_n})^{\frac{1}{n}} = \lim_{n\to+\infty} (\mathbb{E}w(\theta)^{L_n})^{\frac{1}{n}} = h(\theta),$$

where h is defined in the theorem. It is also straigtforward to check that h is essentially smooth. Thus, Gartner-Ellis theorem implies that the laws for $\{\frac{T_n}{n}\}$ satisfies the large deviations principle with rate function

$$I(x) = \sup_\theta [\theta x - \log h(\theta)].$$

□

We end this section with two remarks.

Remark 3.8. *As an alternative to Gartner-Ellis theorem, one can compute the rate function for $\frac{T_n}{n}$ directly as follows:*

$$\mathbb{P}(\frac{T_n}{n} \in (\alpha - \epsilon, \alpha + \epsilon)) = \sum_\beta \mathbb{P}(\frac{T_n}{n} \in (\alpha - \epsilon, \alpha + \epsilon) | \frac{L_n}{n} \in (\beta - \delta, \beta + \delta)) \cdot \mathbb{P}(\frac{L_n}{n} \in (\beta - \delta, \beta + \delta))$$

where the sum is taken over appropriate β's $\in (0,1)$. In each product, the first probability is known as $T_n|L_n$ has a binomial distribution, while the second probability can be aysmptotically computed from the LDP for L_n. Finally, sending $\delta \to 0$, one can get the rate function for $\{\frac{T_n}{n}\}$.

Remark 3.9. *We give a heuristic explaination why the rate function of $\{\frac{L_n}{n}\}$ takes this form. Let $w = e^\eta$, and consider the Laplace transform*

$$(\mathbb{E}w^{L_n})^{\frac{1}{n}} = (\mathbb{E}e^{\eta L_n})^{\frac{1}{n}}.$$

For $w > w^* = \frac{1}{\sqrt{2d-1}}$, the limit exists and is equal to $\frac{2d-1}{2d}e^\eta + \frac{1}{2d}e^{-\eta}$, where $w = e^\eta$. Call the limit of the right hand side $\Lambda(\eta)$, if it exists. One sees that $\eta^* = \log w^*$ is the minimizer of Λ, and thus

$$\Lambda(\eta) \geq \frac{\sqrt{2d-1}}{d}$$

for all $\eta < \eta^*$. On the other hand, if $\Lambda(\eta)$ exists, it must be convex. Because $L_n \geq 0$, it also must not (strictly) increase when η becomes smaller. So the convexity in η forces the graph to the left of $\eta^* = \log w^*$ to be flat, and is thus equal to $\frac{\sqrt{2d-1}}{d}$.

4 Central limit theorem and the invariance principle

In this section, we derive asymptotic distribution of T_n for large n, including central limit theorem and the invariance principle. Under the natural coupling $L_n = S_n + D_n$, $D_n \le 2R_n$, where R_n is (almost) surely bounded by a geometric random variable R. As we will be frequently using this property, we state it as a lemma below.

Lemma 4.1. *Let $L_n = S_n + D_n$, and R_n be defined as above. Then, $D_n \le wR_n$, and $R = \sup_n R_n$ has a geometric distribution with $\mathbb{P}(R_n = k) = \frac{2d-2}{2d-1} \cdot (\frac{1}{2d-1})^{\frac{1}{n}}$.*

Theorem 4.2. *Under the above assumptions, $\frac{1}{\sqrt{n}}(L_n - \frac{d-1}{d}n)$ converges in distribution to $N(0, \frac{2d-1}{d^2})$, and $\frac{1}{\sqrt{n}}(T_n - \frac{2(d-1)^2}{d(2d-1)}n)$ converges in distribution to $N(0, \frac{2(d-1)^2(5d-2)}{d^2(2d-1)^2})$.*

Proof. For any $x \in \mathbb{R}$,

$$\begin{aligned}
F_n(x) &= \mathbb{P}(\frac{1}{\sqrt{n}}(L_n - \frac{2d-1}{d^2}n) \le x) \\
&= \mathbb{P}(\frac{1}{\sqrt{n}}(S_n - \frac{2d-1}{d^2}n) + \frac{D_n}{\sqrt{n}} \le x) \\
&\to F(x),
\end{aligned}$$

where F is the distribution function for $N(0, \frac{2d-1}{d^2})$. The convergence in the last line follows from the fact that $\frac{D_n}{\sqrt{n}} \to 0$ almost surely, as $R = \sup_n R_n$ is almost surely finite.

Now we compute the moment generating function for $\frac{1}{\sqrt{n}}(T_n - \frac{2(d-1)^2}{d(2d-1)}n)$. For simplicity, let $\mu = \frac{2(d-1)^2}{d(2d-1)}$ and $p = \frac{2d-2}{2d-1}$.

$$\begin{aligned}
\mathbb{E}e^{\frac{\theta}{\sqrt{n}}(T_n - \mu n)} &= e^{-\theta\mu\sqrt{n}}\mathbb{E}e^{\frac{\theta}{\sqrt{n}}T_n} \\
&= e^{-\theta\mu\sqrt{n}}\mathbb{E}\mathbb{E}(e^{\frac{\theta}{\sqrt{n}}T_n}|L_n) \\
&= e^{-\theta\mu\sqrt{n}}\mathbb{E}(1 - p + pe^{\frac{\theta}{\sqrt{n}}})^{L_n - 1} \\
&= e^{-\theta\mu\sqrt{n}}\mathbb{E}e^{(L_n-1)\log(1+\frac{p\theta}{\sqrt{n}}+\frac{p\theta^2}{2n}+o(\frac{1}{n}))} \\
&= e^{-\theta\mu\sqrt{n}}\mathbb{E}e^{(L_n-1)[\frac{p\theta}{\sqrt{n}}+\frac{1}{2n}(p\theta^2-p^2\theta^2)+o(\frac{1}{n})]} \\
&= \mathbb{E}e^{\frac{p\theta}{\sqrt{n}}(L_n-\frac{\mu}{p}n)+\frac{1}{2}\theta^2 p(1-p)\frac{L_n}{n}+o(1)} \\
&\to e^{\frac{(d-1)^2(5d-2)}{d^2(2d-1)^2}\theta^2}
\end{aligned}$$

The convergence in the last line follows from the law of large numbers and central limit theorem for L_n, and dominated convergence. This implies that $\frac{1}{\sqrt{n}}(T_n - \frac{2(d-1)^2}{d(2d-1)}n)$ converges in distribution to $N(0, \frac{2(d-1)^2(5d-2)}{d^2(2d-1)^2})$.

$\square$

Since L_n behaves very much like S_n, one expects that under proper scaling, it converges to the standard Brownian motion.

Proposition 4.3. *Let $(\Omega, \mathbb{P}, \mathcal{F})$ be a probability space affording the discrete time process $\{S_i, i \in \mathbb{N}\}$. For each $n \in \mathbb{N}$, define a $C^0([0,1])$-valued random variable $\{W_t^{(n)} : t \in [0,1]\}$ by*

$$W_t^{(n)} = [L_{tn} - \frac{d-1}{d}tn]/\sqrt{n\frac{2d-1}{d^2}}$$

for $t \in \frac{1}{n}[n]$, and linearly interpolated for other values of t. Then the sequence converges in law to the standard one dimensional Brownian motion on $[0,1]$ as $n \to \infty$.

Proof. This is plain in light of Lemma 3.4, as $L_{[tn]} - S_{[tn]}$ is bounded by a geometric random variable, uniformly in t. $\square$

More difficult is the invariance principle for T_n, the number of turns at time n. We prove it in the next theorem. Let $\sigma^2 = \frac{2(d-1)^2(5d-2)}{d^2(2d-1)^2}$, then,

Theorem 4.4. *For each n,define a $C^0([0,1])$-valued random variable $\{W_t^{(n)} : t \in [0,1]\}$ by*

$$W_t^{(n)} = \frac{1}{\sigma\sqrt{n}}[T_{tn} - \frac{2(d-1)^2}{d(2d-1)}tn] \tag{7}$$

for $t \in \frac{1}{n}[n]$, and linearly interpolated for other values of t. Then the sequence converges in law to the standard one dimensional Brownian motion on $[0,1]$ as $n \to \infty$.

Proof. Recall for a sequence of processes to converge to Brownian motion, it suffices to check that their finite dimensional joint distributions converge to that of a Brownian motion and Prohorov tightness criterion (see Theorem 16.5 of [5]). We prove them in the two lemmas below.

Lemma 4.5. *Let $W_t^{(n)}$ be defined as above. Then,*

$$(W_{t_1}^{(n)}, W_{t_2}^{(n)}, \cdots, W_{t_k}^{(n)}) \to (B_{t_1}, B_{t_2}, \cdots, B_{t_k}),$$

where B_t is a standard one-dimensional Brownian motion starting at 0.

Proof. It suffices to show that $\mathcal{L}(W_t^{(n)} - W_s^{(n)} | W_r^{(n)} : r \le s)$ is asymptotically $N(0, t-s)$ and independent of $\{W_r^{(n)} : r \le s\}$. This is a generalization of the central limit theorem for T_n, which corresponds to the case $s = 0$, and also suggests that for general s, we can compare with the case $s = 0$. We couple $W_t^{(n)} - W_s^{(n)}$ with two extremely cases, as considered below.

Let $T_{(\cdot)}$ be the standard process of number of turns, and define

$$U_{(m,n)} := T|_{[m,n]},$$

that is, the number of turns in the segment of the walk during the time interval $[m, n]$. Then, one immediately sees that

$$T_{\lfloor tn \rfloor} - T_{\lfloor sn \rfloor} \le U_{[\lfloor sn \rfloor, \lfloor tn \rfloor]},$$

as the former may cancel turns created before time sn.

On the other hand, we have the reversed inequality that

$$T_{\lfloor tn \rfloor} - T_{\lfloor sn \rfloor} \ge U_{[\lfloor sn \rfloor, \lfloor tn \rfloor]} - 2|\min_{1 \le k \le \lfloor (t-s)n \rfloor} S_k|,$$

and it suffices to estimate the behavior of $|\min_{1 \le k \le n} S_k|$. Since

$$\mathbb{P}(|\min_{1 \le k \le n}| = M) \le \sum_{k=0}^{\lfloor (t-s)n \rfloor} \mathbb{P}(S_k = -M) \le C(\frac{\sqrt{2d-1}}{d})^M,$$

we see that it is bounded by a geometric random variable, and thus

$$\frac{1}{\sigma n}|\min_{1 \le k \le (t-s)n} S_k| \to 0$$

in probability. Note that $U_{[m,n]}$ has the same distribution as T_{n-m}, therefore $\mathcal{L}(W_t^n - W_s^n | W^n(0,s))$ converges to $N(0, t-s)$, and is clearly independent of $W^n(0,2)$. By induction, we also get the asymptitically independent increment property for the sequence W^n.

$\square$

Lemma 4.6. $W_t^{(n)}$ *satisfies Prohorov's tightness condition, i.e.,*

$$\lim_{h\to 0}\limsup_{n\to\infty}\mathbb{E}[\sup_{|t-s|\leq h}|W_t^{(n)}-W_s^{(n)}|\wedge 1]=0$$

Proof. Let $\bar{T}_n := T_n - \frac{2(d-1)^2}{d(2d-1)}n$ be the centered number of turns at step n. The idea of the proof is similar to Ottaviani's maximal inequality for random walk on $\mathbb{R}$ (see Lemma 14.8 of [5]).

Fix $\epsilon > 0$. Let $t \in [0,1)$ and $h \in (0, 1-t)$. Define $\tau := \min\{k \in (tn, (t+h)n] : |\bar{T}_k - \bar{T}_{[tn]}| \geq 2\epsilon\sigma\sqrt{n}\}$. Then,

$$\begin{aligned}\mathbb{P}(|\bar{T}_{\lfloor(t+h)n\rfloor}-\bar{T}_{\lfloor tn\rfloor}|>\epsilon\sigma\sqrt{n}) &\geq \mathbb{P}(\tau\leq(t+h)n, |\bar{T}_{\lfloor(t+h)n\rfloor}-\bar{T}_{\tau}|\leq\epsilon\sigma\sqrt{n})\\ &= \sum_{k=\lfloor tn\rfloor+1}^{\lfloor(t+h)n\rfloor}\mathbb{P}(\tau=k)\mathbb{P}(|\bar{T}_{\lfloor(t+h)n\rfloor}-\bar{T}_k|\leq\epsilon\sigma\sqrt{n}|\tau=k)\\ &\geq \mathbb{P}(\tau\leq n)\min_{k\in(tn,(t+h)n]}\mathbb{P}(|\bar{T}_{\lfloor(t+h)n\rfloor}-\bar{T}_k|\leq\epsilon\sigma\sqrt{n}|\tau=k).\end{aligned}$$

We want to bound $\mathbb{P}(\tau \leq n) = \mathbb{P}(\max_k |\bar{T}_k - \bar{T}_{\lfloor tn\rfloor}| > 2\epsilon\sigma\sqrt{n})$. First note that when hn is large enough, $\bar{T}_{\lfloor(t+h)n\rfloor} - \bar{T}_{\lfloor tn\rfloor}$ behaves like a Gaussian with mean 0 and variance $h\sigma^2 n$. So, there exists $N(\epsilon)$ such that for all $n > \frac{N(\epsilon)}{h}$, we have

$$\mathbb{P}(|\bar{T}_{\lfloor(t+h)n\rfloor}-\bar{T}_{\lfloor tn\rfloor}|>\epsilon\sigma\sqrt{n})\leq\exp(-\epsilon^2/2h). \tag{8}$$

On the other hand, $|\bar{T}_{\lfloor(t+h)n)\rfloor} - \bar{T}_k| \leq U_{(k,(t+h)n)} + M_{\lfloor hn\rfloor}$. According to Proposition (2.5), the variance of the former term on the right hand side is to $(t+h)n - k$ with an error uniformly bounded in t, h, k and n. The second term is dominated by a geometric random variable, and thus has a finite variance. So, by Chebyshev's inequality, we have

$$\mathbb{P}(|\bar{T}_{\lfloor(t+h)n\rfloor}-\bar{T}_k|\leq\epsilon\sigma\sqrt{n}|\tau=k)\geq 1-\frac{2[(t+h)n-k+C]}{n}.$$

Since k takes values between tn and $(t+h)n$, the minimum bound is achieved at $k = \lfloor tn\rfloor + 1$, so we get

$$\min_{k\in(tn,(t+h)n]}\mathbb{P}(|\bar{T}_{\lfloor(t+h)n\rfloor}-\bar{T}_k|\leq\epsilon\sigma\sqrt{n}|\tau=k)\geq 1-\frac{2h}{\epsilon^2}-\frac{C}{n},$$

where C is independent of t, h, k and n. The last inequality is valid as long as the right hand side is positive, which requires h to take small values and n to take large values. These values depend on ϵ only. The last bound, together with the bound (8), imply

$$\mathbb{P}(\max_{k\in(tn,(t+hn)]}|\bar{T}_k-\bar{T}_{tn}|>2\epsilon\sigma\sqrt{n})\leq(1-\frac{2h}{\epsilon^2}-\frac{C}{n})^{-1}\exp(-\epsilon^2/2h). \tag{9}$$

Since for $n > \frac{16}{\epsilon^2\sigma^2}$ we have

$$\mathbb{P}(\sup_{\delta\in(0,h)}|W_{t+\delta}^{(n)}-W_t^{(n)}|>\epsilon)\leq\mathbb{P}(\sup_{k\in(tn,(t+h)n)}|\bar{T}_k-\bar{T}_{\lfloor tn\rfloor}|>\frac{\epsilon}{2}\sigma\sqrt{n}),$$

and the bound on the right hand side of (9) is independent of t, one gets

$$\sup_{t\in(0,1)}\mathbb{P}(\sup_{\delta\in(0,h)}|W_{t+\delta}^{(n)}-W_t^{(n)}|>\epsilon)<(1-\frac{2h}{\epsilon^2}-\frac{C}{n})^{-1}\exp(-\epsilon^2/32h)$$

for all $n > \max\{\frac{N(\epsilon)}{h}, \frac{16}{\epsilon^2\sigma^2}\}$. Now divide the interval $(0,1)$ into $\lfloor\frac{1}{h}\rfloor + 1$ subintervals, each with length at most h. Then, $|t-s| < h$ implies that either s, t are in the same subinterval, or they are in two adjacent ones. This observation gives

$$\mathbb{P}(\sup_{|t-s|<h}|W_t^{(n)}-W_s^{(n)}|>\epsilon)<(\frac{2}{h}+2)(1-\frac{2h}{\epsilon^2}-\frac{C}{n})^{-1}\exp(-\epsilon^2/32h) \tag{10}$$

for all $n > \max\{N(\epsilon), \frac{16}{\epsilon^2\sigma^2}\}$. Since

$$\mathbb{E}[\sup_{|t-s|<h}|W_t^{(n)}-W_s^{(n)}|\wedge 1]\leq\epsilon+\mathbb{P}(\sup_{|t-s|<h}|W_t^{(n)}-W_s^{(n)}|>\epsilon),$$

the maximal inequality (10) quickly gives

$$\lim_{h\to 0}\limsup_{n\to\infty}\mathbb{E}[\sup_{|t-s|\le h}|W_t^{(n)}-W_s^{(n)}|\wedge 1]\le\epsilon,$$

which implies Prohov's tightness condition since ϵ is arbitrary.

□

Combining the above two lemmas, we prove the invariance principle.

□

References

[1] K.T.Chen (1957), Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula, *Annals of Mathematics*, 65(1), pp.163-178.

[2] A.Dembo, O.Zeitouni (1998), Large Deviations Techniques and Applications, *Stochastic Modelling and Applied Probability.*

[3] W.Feller (1968), An Introduction to Probability Theory and its Applications, Vol.1, *Wiley series.*

[4] B.M.Hambly, T.J.Lyons (2010), Uniqueness for the signature of a path of bounded variation and the reduced path group, *Annals of Mathematics*, 171(1), 109-167.

[5] O.Kallenberg (2002), Foundations of Modern Probability, second edition, *Springer.*

[6] T.J.Lyons, W.Xu (2011), Inversion of signature for paths of bounded variation, in preparation.

[7] S.Orey (1958), A central limit theorem for m-dependent random variables, *Duke Math Journal*, 25(4), 543-546.